Center for Advanced Computation

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
URBANA, ILLINOIS 61801

CAC Document No. 57

ON JACOBI AND JACOBI-LIKE ALGORITHMS FOR A PARALLEL COMPUTER

By

Ahmed H. Sameh

December 11, 1972

CAC Document No. 57

ON JACOBI AND JACOBI-LIKE ALGORITHMS FOR A PARALLEL COMPUTER

By

Ahmed H. Sameh

Center for Advanced Computation
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

December 11, 1972

This work was supported in part by the Advanced Research Projects Agency of the Department of Defense, was monitored by the U. S. Army Research Office-Durham under Contract No. DAHC04-72-C-0001, and appeared in Mathematics of Computation, July, 1971, Vol. 25, No. 115.

ABSTRACT

Many existing algorithms for obtaining the eigenvalues and eigenvectors of matrices would make poor use of such a powerful parallel computer as the ILLIAC IV. In this paper Jacobi's algorithm for real symmetric or complex Hermitian matrices, and a Jacobi-like algorithm for real non-symmetric matrices developed by P. J. Eberlein, are modified so as to achieve maximum efficiency for the parallel computations.

TABLE OF CONTENTS

1. Introduction. With the advent of parallel computers, the study of computationally massive problems became economically possible. Such problems include, for example, solution of sets of partial differential equations over sizable grids, and multiplication, inversion, or determination of eigenvalues and eigenvectors of large matrices.

An example of a parallel computer is the ILLIAC IV . This computer is essentially an array of coupled arithmetic units driven by instructions from a common control unit. Each of the arithmetic units, called processing elements (PE's), have 2048 words of 64-bit memory with an access time under 420 nanoseconds. Each PE is capable of 64-bit floating point multiplication in about 550 nanoseconds. Two 32-bit floating point operations may be performed in each PE in approximately the same times. The PE instruction set is similar to that of conventional machines with two exceptions. First, the PE's are capable of communicating data to four neighboring PE's by means of routing instructions. Second, the PE's are able to set their own mode registers to effectively disable or enable themselves. For more detailed description of this system the reader is referred to [2, 8, 9, 12].

The purpose of this paper is to introduce modified Jacobi and Jacobi-like algorithms for the computation of the eigenvalues and eigenvectors of large real symmetric or complex Hermitian matrices, and real non-symmetric matrices respectively, that are suitable for a parallel computer.

2. Jacobi's Algorithm. In the classical method of Jacobi (1846), [13], a real symmetric matrix is reduced to the diagonal form by a sequence of plane rotations $A_{k+1} = R_k A_k R_k^t$ $(k = 1,2,\ldots.)$, where $A_1 = A$ is the original matrix and each rotation $R_k \equiv R(p,q,\alpha_{pq}^{(k)})$ in the p, q plane through an angle $\alpha_{pq}^{(k)}$ eliminates the off-diagonal element $a_{pq}^{(k)}$ (and hence $a_{qp}^{(k)}$), and affects only elements in rows and columns p and q. (See Appendix for the appropriate value of $\alpha_{pq}^{(k)}$ to annihilate the element $a_{pq}^{(k)}$.) Because of symmetry only the off-diagonal elements above the main diagonal are considered in what follows.

It is possible, however, to modify the present Jacobi algorithm for a parallel computer so as to eliminate more than one off-diagonal element. For example, for a matrix A of order 4, if the orthogonal transformation R is chosen as,

$$(2.1) \qquad R = \begin{bmatrix} c_1 & 0 & s_1 & 0 \\ 0 & c_2 & 0 & s_2 \\ -s_1 & 0 & c_1 & 0 \\ 0 & -s_2 & 0 & c_2 \end{bmatrix},$$

where $c_i = \cos \alpha_i$, $s_i = \sin \alpha_i$ $(i = 1,2)$, then $R\,A\,R^t$ would have zero elements in positions (1,3) and (2,4) provided that the angles α_1 and α_2 are properly chosen. α_1 and α_2 are determined by (a_{11}, a_{33}, a_{13}) and (a_{22}, a_{44}, a_{24}) respectively.

Define m by $[(n + 1)/2]$, where n is the order of the matrix A and $[x]$ is the greatest integer less than or equal to x. Let each (2m - 1) orthogonal transformations be denoted by a sweep. Observing that there are n(n - 1)/2 off-diagonal elements, and that the maximum number of these elements which can be annihilated by an orthogonal transformation of the type (2.1) is [n/2], then the modified Jacobi algorithm will attain maximum efficiency of parallel computation if the following two conditions are satisfied:

(i) each orthogonal transformation R_k should be constructed so as to annihilate [n/2] off-diagonal elements.

(ii) each sweep should annihilate each off-diagonal element only once; i.e., each of the (2m - 1) orthogonal transformations in a sweep should annihilate different [n/2] off-diagonal elements.

Several annihilation regimes that satisfy the above requirements are possible. Two different regimes are discussed below.

First Annihilation Regime. For a given sweep each of the (2m - 1) orthogonal matrices R_k consists of the elements,

$$R^{(k)}_{pp} = R^{(k)}_{qq} = \cos \alpha^{(k)}_{pq}; \; R^{(k)}_{pq} = -R^{(k)}_{qp} = \begin{cases} \sin \alpha^{(k)}_{pq} & p < q, \\ -\sin \alpha^{(k)}_{pq} & p > q, \end{cases} \tag{2.2}$$

where p and q are sequences defined by

(a) for $k = 1,2,\ldots, m - 1$,

$q = m - k + 1, m - k + 2, \ldots, n - k$,

$$p = \begin{cases} (2m - 2k + 1) - q, & m - k + 1 \le q \le 2m - 2k, \\ (4m - 2k) - q, & 2m - 2k < q \le 2m - k - 1, \\ n, & 2m - k - 1 < q, \end{cases} \tag{2.3}$$

(b) for $k = m, m + 1, \ldots, 2m - 1$,

$q = 4m - n - k, 4m - n - k + 1, \ldots, 3m - k - 1$,

$$p = \begin{cases} n, & q < 2m - k + 1, \\ (4m - 2k)-q, & 2m - k + 1 \le q \le 4m - 2k - 1, \\ (6m - 2k - 1)-q, & 4m - 2k - 1 < q. \end{cases} \tag{2.4}$$

The remaining elements of R_k are zero except for n odd, then $R^{(k)}_{2m-k,2m-k} = 1$. For a given k, the angles $\alpha^{(k)}_{pq}$ are determined for all (p,q) such that $\alpha^{(k)}_{pq}$ eliminates the element $a^{(k)}_{pq}$; see Appendix.

Let n = 8 and k = 2, then the pairs (p,q) are given by {(2,3); (1,4); (7,5); (8,6)} and R_2 is of the form,

$$
\begin{bmatrix}
R^{(2)}_{11} & & & R^{(2)}_{14} & & & & \\
 & R^{(2)}_{22} & R^{(2)}_{23} & & & & & \\
 & -R_{23} & R^{(2)}_{33} & & & & & \\
-R^{(2)}_{14} & & & R^{(2)}_{44} & & & & \\
 & & & & R^{(2)}_{55} & & R^{(2)}_{57} & \\
 & & & & & R^{(2)}_{66} & & R^{(2)}_{68} \\
 & & & & -R^{(2)}_{57} & & R^{(2)}_{77} & \\
 & & & & & -R^{(2)}_{68} & & R^{(2)}_{88}
\end{bmatrix}
$$

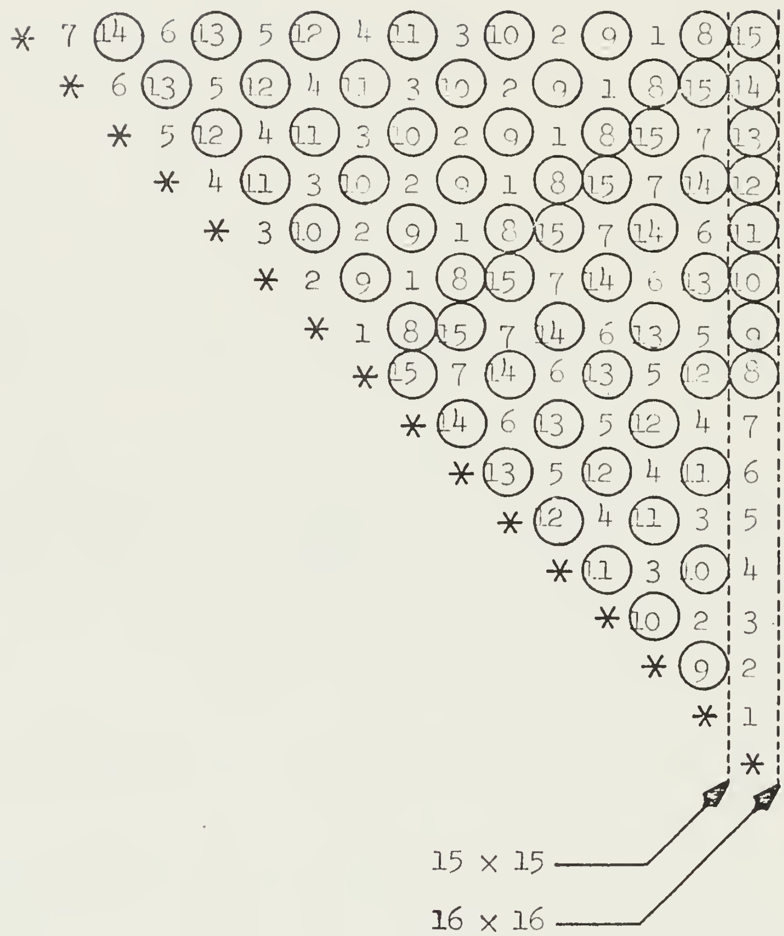
while for k = 7 the pairs (p,q) are {(8,1); (7,2); (6,3); (5,4)} and R_7 is of the form

$$
\begin{bmatrix}
R^{(7)}_{11} & & & & & & & R^{(7)}_{18} \\
 & R^{(7)}_{22} & & & & & R^{(7)}_{27} & \\
 & & R^{(7)}_{33} & & & R^{(7)}_{36} & & \\
 & & & R^{(7)}_{44} & R^{(7)}_{45} & & & \\
 & & & -R^{(7)}_{45} & R^{(7)}_{55} & & & \\
 & & -R^{(7)}_{36} & & & R^{(7)}_{66} & & \\
 & -R^{(7)}_{27} & & & & & R^{(7)}_{77} & \\
-R^{(7)}_{18} & & & & & & & R^{(7)}_{88}
\end{bmatrix}
$$

If the order of the matrix is odd, say $n = 7$, then for $k = 3$ the pairs (p,q) are given by $\{(1,2);\ (7,3);\ (6,4)\}$ and R_3 is of the form,

$$\begin{bmatrix} R_{11}^{(3)} & R_{12}^{(3)} & & & & & \\ -R_{12}^{(3)} & R_{22}^{(3)} & & & & & \\ & & R_{33}^{(3)} & \cdots & \cdots & \cdots & R_{37}^{(3)} \\ & & \vdots & R_{44}^{(3)} & \cdots & R_{46}^{(3)} & \vdots \\ & & \vdots & \vdots & 1 & \vdots & \vdots \\ & & \vdots & -R_{46}^{(3)} & \cdots & R_{66}^{(3)} & \vdots \\ & & -R_{37}^{(3)} & \cdots & \cdots & \cdots & R_{77}^{(3)} \end{bmatrix}$$

For example, in a given sweep, denoting each element eliminated in the k-th transformation by the integer k, the patterns of the annihilated elements for matrices of orders 16 and 15 are shown below.


15 × 15
16 × 16

<u>Second Annihilation Regime</u>. This regime satisfies conditions (i) and (ii) for matrices of order $n = 2^{\gamma}$, where γ is an integer. The elements of each orthogonal transformation, in a given sweep, R_k $(k = 1,2,\ldots, n-1)$ are given by (2.2). For $k = 1,2,\ldots, n/2$ the pairs (p,q) are defined by,

$$q = 2,4,6, \ldots, n,$$

$$p = \begin{cases} q + (n - 2k + 1), & q < 2k , \\ q - 2k + 1, & q \geq 2k . \end{cases} \tag{2.5}$$

Let $n = 8$ and $k = 3$, then the pairs (p,q) are $\{(5,2); (7,4); (1,6); (3,8)\}$ and R_3 is of the form

$$\begin{bmatrix}
R_{11}^{(3)} & & & & & R_{16}^{(3)} & & \\
& R_{22}^{(3)} & & & R_{25}^{(3)} & & & \\
& & R_{33}^{(3)} & & & & & R_{38}^{(3)} \\
& & & R_{44}^{(3)} & & & R_{47}^{(3)} & \\
& -R_{25}^{(3)} & & & R_{55}^{(3)} & & & \\
-R_{16}^{(3)} & & & & & R_{66}^{(3)} & & \\
& & & -R_{47}^{(3)} & & & R_{77}^{(3)} & \\
& & -R_{38}^{(3)} & & & & & R_{88}^{(3)}
\end{bmatrix}$$

In order to construct the orthogonal transformations R_k for $k = n/2 + 1,\ n/2 + 2, \ldots, n-1$; consider the sequence $L = 1,2,\ldots, \gamma - 1$. For each value of L there are $N = 2^{\gamma-L-1}$ orthogonal matrices R_k given by,

$$R_k = \text{diag}\ (H_1^{(k)},\ H_2^{(k)}, \ldots,\ H_t^{(k)}) \tag{2.6}$$

where $t = 2^{L-1}$, $k = n(1-2^{-L}) + \ell$, and $\ell = 1,2,\ldots, N$. The sequences p and q for each $H_M^{(k)}$, $(M = 1,2,\ldots,t)$, are defined by,

$$p = i + 4N\,(M-1)\ , \qquad\qquad i = 1,2,\ldots,2N\ ,$$

$$q = p + 2(N+\ell-1) - 2N[O(1)]\ ,$$

where,

$$O(1) = \begin{cases} 0\ , & i + 2(N+\ell-1) \leq 4N\ , \\ 1\ , & \text{otherwise}\ . \end{cases}$$

Let $n = 8$, $L = 2$, and $\ell = 1$, then $k = 7$, and the pairs (p,q) are given by $\{(1,3);\ (2,4);\ (5,7);\ (6,8)\}$ and R_7 is of the form,

$$\begin{bmatrix} R_{11}^{(7)} & & R_{13}^{(7)} & & & & & \\ & R_{22}^{(7)} & & R_{24}^{(7)} & & & & \\ -R_{13}^{(7)} & & R_{33}^{(7)} & & & & & \\ & -R_{24}^{(7)} & & R_{44}^{(7)} & & & & \\ & & & & R_{55}^{(7)} & & R_{57}^{(7)} & \\ & & & & & R_{66}^{(7)} & & R_{68}^{(7)} \\ & & & & -R_{57}^{(7)} & & R_{77}^{(7)} & \\ & & & & & -R_{68}^{(7)} & & R_{88}^{(7)} \end{bmatrix}$$

The pattern of the annihilated elements in one sweep for a matrix of order 16 is shown below, where those elements annihilated in the k-th transformation are denoted by the integer k.

*	1	⑮	2	⑬	3	⑭	4	⑨	5	⑩	6	⑪	7	⑫	8
	*	8	⑮	7	⑬	6	⑭	5	⑨	4	⑩	3	⑪	2	⑫
		*	1	⑭	2	⑬	3	⑫	4	⑨	5	⑩	6	⑪	7
			*	8	⑭	7	⑬	6	⑫	5	⑨	4	⑩	3	⑪
				*	1	⑮	2	⑪	3	⑫	4	⑨	5	⑩	6
					*	8	⑮	7	⑪	6	⑫	5	⑨	4	⑩
						*	1	⑩	2	⑪	3	⑫	4	⑨	5
							*	8	⑩	7	⑪	6	⑫	5	⑨
								*	1	⑮	2	⑬	3	⑭	4
									*	8	⑮	7	⑬	6	⑭
										*	1	⑭	2	⑬	3
											*	8	⑭	7	⑬
												*	1	⑮	2
													*	8	⑮
														*	1
															*

Using one quadrant of the ILLIAC IV, (64 PE's), then for a 128 × 128 matrix, the 64 angles of each transformation are determined simultaneously, one angle per PE. Once the transformation matrix R_k is determined, the matrix $A_{k+1} = R_k A_k R_k^t$ is computed in parallel [7]. Assuming that the matrix has converged, (using some criterion [13]), to the diagonal form after u sweeps, or after $r - 1 = u\ (2m-1)$ orthogonal transformations, then the diagonal elements of $A_r = WAW^t$ are taken to be the eigenvalues of A. The columns of $W^t = (V_u V_{u-1} \cdots \cdot V_1)^t$ are the corresponding eigenvectors, where for the j-th sweep $V_j = \prod_{k=1}^{2m-1} (R_k)_j$, $(j = 1,2,\ldots,u)$.

A similar algorithm as that described above [11] has been programmed in ILLIAC IV assembly language and successfully tested on an ILLIAC IV execution simulator [1].

3. A Jacobi-Like Algorithm for Nonsymmetric Matrices.

Eberlein [3,4] showed that for an $n \times n$ matrix A, complex in general, there exists a matrix $U = \Pi_\ell U_\ell (k,m)$ generated from a sequence of two dimensional transformations $U_\ell (k,m)$, where (k,m) is the pivot pair, such that $A_L = U^{-1} A U$ is arbitrarily close to being normal; i.e., the matrix $(A_L A_L^* - A_L^* A_L)$ is arbitrarily small. At each stage of the iteration, based on the elements of the k-th and m-th rows and columns, the parameters of U_ℓ were chosen such that the decrement of the Euclidean norm of A_ℓ is given by,

$$N^2(A_\ell) - N^2(U_\ell^{-1} A_\ell U_\ell) \geq [1/3n(n-1)].\ N^2(A_\ell A_\ell^* - A_\ell^* A_\ell)$$

where, $N^2(A) = \sum_{i,j} |a_{ij}|^2$.

In this paper, the above algorithm has been modified for parallel computation. The transformations U_ℓ are n-dimensional, and their parameters are based on all the elements of the matrix A_ℓ. A lower bound on the decrement of the Euclidean norm of A_ℓ is given by,

$$N^2(A_\ell) - N^2(U_\ell^{-1} A_\ell U_\ell) \geq (1/4n)\ N^2(A_\ell A_\ell^* - A_\ell^* A_\ell).$$

Once the matrix is practically normal, one can use the optimal procedure of Goldstine and Horwitz [5] for reducing it to the diagonal form, thus the eigenvalues and eigenvectors of A are obtained.

Since a nondiagonable matrix cannot be similar to a normal matrix, then this procedure yields its best results for diagonable matrices, (see example 7 in [3], p. 84).

Let the original matrix A be real, diagonable, and of an even order $n = 2r$, (if n is odd A is replaced by diag (A,ν) of order $n + 1$), then it can be partitioned as follows,

$$
(3.1) \qquad A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ A_{21} & A_{22} & \cdots & A_{2r} \\ \vdots & \vdots & & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rr} \end{bmatrix}
$$

where each submatrix A_{km}, $(k, m = 1, 2, \ldots, r)$,

is given by,

$$
(3.2) \qquad A_{km} = \begin{bmatrix} a_{2k-1,\ 2m-1} & a_{2k-1,\ 2m} \\ a_{2k,\ 2m-1} & a_{2k,\ 2m} \end{bmatrix}.
$$

Let,

$$
(3.3) \qquad \begin{aligned} D_{km} &= (a_{2k-1,\ 2m-1} - a_{2k,\ 2m}) \\ E_{km} &= (a_{2k-1,\ 2m} - a_{2k,\ 2m-1}) \\ B_{km} &= (a_{2k-1,\ 2m} + a_{2k,\ 2m-1}) \end{aligned}
$$

and,

$$(3.4) \qquad \begin{aligned} \kappa_1(A) &= \sum_{k,m} (D_{km}^2 + E_{km}^2) \\ \kappa_2(A) &= \sum_{k,m} D_{km} E_{km} \end{aligned}$$

Assume also that A has been scaled such that $N^2(A) \leq 1$, and denote the matrix $(A A^t - A^t A)$ by C.

Lemma 1. Let $A' = Q^{-1} A Q$, where $Q = \text{diag}(S_1, S_2, \ldots, S_r)$, and $S_1 = S_2 = \ldots = S_r = S$ is given by,

$$(3.5) \qquad S = \begin{bmatrix} \cosh \Psi & \sinh \Psi \\ \sinh \Psi & \cosh \Psi \end{bmatrix}.$$

Define Ψ by,

$$(3.6) \qquad \tanh 4\Psi = -2\kappa_2(A)/\kappa_1(A).$$

Provided that $\kappa_1(A) > 2\,|\kappa_2(A)|$, the following relation holds,

$$(3.7) \qquad \Delta N^2(A) \geq \kappa_2^2(A)/\kappa_1(A)$$

where $\Delta N^2(A) = N^2(A) - N^2(A')$ is the decrement of the Euclidean norm of A.

Proof. The elements of each submatrix $A'_{km} = S^{-1} A_{km} S$ are given by

$$a'_{2k-1,2m-1} = a_{2k-1,2m-1} \cosh^2 \psi - a_{2k,2m} \sinh^2 \psi + \frac{1}{2} E_{km} \sinh 2\psi$$

$$a'_{2k,2m} = -a_{2k-1,2m-1} \sinh^2 \psi + a_{2k,2m} \cosh^2 \psi - \frac{1}{2} E_{km} \sinh 2\psi$$

(3.8) $$a'_{2k-1,2m} = \frac{1}{2} D_{km} \sinh 2\psi + a_{2k-1,2m} \cosh^2 \psi - a_{2k,2m-1} \sinh^2 \psi$$

$$a'_{2k,2m-1} = -\frac{1}{2} D_{km} \sinh 2\psi - a_{2k-1,2m} \sinh^2 \psi + a_{2k,2m-1} \cosh^2 \psi$$

Therefore,

$$N^2 (A'_{km}) = N^2 (A_{km}) + (D^2_{km} + E^2_{km}) \sinh^2 2\psi + D_{km} E_{km} \sinh 4\psi$$

and consequently,

(3.9) $$\Delta N^2 (A_{km}) = -D_{km} E_{km} \sinh 4\psi - \frac{1}{2} (D^2_{km} + E^2_{km}) (\cosh 4\psi - 1)$$

Since $N^2 (A) = \sum_{k,m} N^2 (A_{km})$, then

(3.10) $$\Delta N^2 (A) = -\frac{1}{2} (\cosh 4\psi - 1) \quad \kappa_1(A) - (\sinh 4\psi) \quad \kappa_2(A).$$

A necessary condition for $\Delta N^2 (A)$ to be an extremum with respect to ψ is $\frac{\partial}{\partial \psi} \Delta N^2 (A) = 0$, this yields relation (3.6),

$$\tanh 4\psi = -2\kappa_2(A)/\kappa_1(A).$$

From the definition (3.4) it is clear that $\kappa_1(A) \geq 2|\kappa_2(A)|$. Excluding for the time being the case $\kappa_1(A) = 2|\kappa_2(A)|$, then the second derivative of $\Delta N^2(A)$ with respect to Ψ evaluated for Ψ in (3.6) is given by,

(3.11) $$-8\kappa_1(A)[1-(4\kappa_2^2(A)/\kappa_1^2(A))] (\cosh 4\psi)$$

and is less than zero. Thus, for the choice (3.6) of ψ, $\Delta N^2(A)$ achieves its maximum value,

$$\Delta N^2(A) = \tfrac{1}{2}\kappa_1(A)[1-\{1-(4\kappa_2^2(A)/\kappa_1^2(A))\}^{\frac{1}{2}}] \tag{3.12}$$

which vanishes only if $\kappa_2(A) = 0$. Since one is considering the case $\kappa_1(A) > 2|\kappa_2(A)|$ then by the binomial theorem,

$$(1 - 4\kappa_2^2(A)/\kappa_1^2(A))^{\frac{1}{2}} = 1 - \tfrac{1}{2}(4\kappa_2^2(A)/\kappa_1^2(A)) - \tfrac{1}{8}(4\kappa_2^2(A)/\kappa_1^2(A))^2 - \ldots \tag{3.13}$$

and (3.12) yields the relation (3.7). If $\kappa_1(A) = 2|\kappa_2(A)|$, then from (3.10), $\Delta N^2(A)$ is given by $\tfrac{1}{2}\kappa_1(A)[1 - \{(1 \pm \tanh 4\psi)/(1 - \tanh^2 4\psi)^{\frac{1}{2}}\}]$. Choosing $\tanh 4\psi = \mp (1 - \epsilon^2)/(1 + \epsilon^2)$, where ϵ is a small number, then $\Delta N^2(A) = \tfrac{1}{2}(1 - \epsilon)\kappa_1(A)$ which is greater than zero.

Lemma 2. Let $A' = P^t A P$, where P is the orthogonal transformation,

$$P = \text{diag}\ (T_1, T_2, \ldots\ldots, T_r) \tag{3.14}$$

in which

$$T_k = \begin{bmatrix} \cos\varphi_k & \sin\varphi_k \\ -\sin\varphi_k & \cos\varphi_k \end{bmatrix} \qquad (k = 1,2,\ldots\ldots,r) \tag{3.15}$$

Then, if φ_k is determined by

$$(3.16) \qquad \tan 2\varphi_k = \frac{c_{2k-1,2k-1} - c_{2k,2k}}{2c_{2k-1,2k}}$$

where c_{ij} are the elements of the matrix $C = AA^t - A^tA$,

$$(3.17) \qquad \kappa_2^2(A') \geq \frac{1}{2n} N^2(C)$$

Proof. The 2×2 diagonal submatrices C_{kk} of the matrix C can be expressed as

$$(3.18) \qquad C_{kk} = \sum_{m=1}^{r} \left[A_{km} A_{km}^t - A_{mk}^t A_{mk}\right] \qquad k = 1,2,\ldots,r$$

Therefore,

$$(3.19) \qquad \sum_{k=1}^{r} C_{kk} = \sum_{k,m=1}^{r} \left[A_{km} A_{km}^t - A_{km}^t A_{km}\right]$$

where,

$$(3.20) \qquad (A_{km} A_{km}^t - A_{km}^t A_{km}) = \begin{bmatrix} E_{km} B_{km} & -D_{km} E_{km} \\ -D_{km} E_{km} & -E_{km} B_{km} \end{bmatrix}$$

Equating the off-diagonal elements of the left and right-hand sides of (3.19),

$$(3.21) \qquad \sum_{k=1}^{r} c_{2k-1,2k} = - \sum_{k,m} D_{km} E_{km} = - \kappa_2(A)$$

Consequently, if the orthogonal matrix P is chosen such that the off-diagonal elements $c_{2k-1,2k}$, for all k, attain their maximum positive values; then the inequality (3.17) is achieved. To show that, consider the matrix $C' = A'A'^t - A'^tA'$. Since $A' = P^t A P$, then $C' = P^t C P$, and the elements of the diagonal submatrices $C'_{kk} = T_k^t C_k T_k$ are given by,

$$
\begin{aligned}
c'_{2k-1,2k} &= c_{2k-1,2k} \cos 2\varphi_k + \frac{1}{2}(c_{2k-1,2k-1} - c_{2k,2k,}) \sin 2\varphi_k \\
c'_{2k-1,2k-1} &= c_{2k-1,2k-1} \cos^2 \varphi_k + c_{2k,2k} \sin^2 \varphi_k - c_{2k-1,2k} \sin 2\varphi_k \\
c'_{2k,2k} &= c_{2k-1,2k-1} \sin^2 \varphi_k + c_{2k,2k} \cos^2 \varphi_k + c_{2k-1,2k} \sin 2\varphi_k
\end{aligned} \tag{3.22}
$$

and

$$c'_{2k,2k-1} = c'_{2k-1,2k}$$

Hence, for $c'_{2k-1,2k}$ to be an extremum (3.16) must hold. Also for the choice (3.16) of φ_k the second derivative of $c'_{2k-1,2k}$ with respect to φ_k is given by,

$$-(h^2/c_{2k-1,2k}) \cos 2\varphi_k \tag{3.23}$$

where, $h = [4c^2_{2k-1,2k} + (c_{2k-1,2k-1} - c_{2k,2k})^2]^{\frac{1}{2}}$. As a result if $\cos 2\varphi_k$ is of the same sign as $c_{2k-1,2k}$, $c'_{2k-1,2k}$ attains its maximum value. Restricting φ_k to the interval $[0,\pi]$, the elements of T_k are given by,

$$
\begin{aligned}
\sin^2 \varphi_k &= \frac{1}{2} - (c_{2k-1,2k}/h) \\
\cos^2 \varphi_k &= \frac{1}{2} + (c_{2k-1,2k}/h)
\end{aligned} \tag{3.24}
$$

in which $\sin \varphi_k > 0$ and $\cos \varphi_k$ is of the same sign as $(c_{2k-1,2k-1} - c_{2k,2k})$. The maximum value of $c'_{2k-1,2k}$ turns out to be $\frac{1}{2}$ h, and

$c'_{2k-1,2k-1} = c'_{2k,2k} = \frac{1}{2}(c_{2k-1,2k-1} + c_{2k,2k})$. Excluding the case when $c_{2k-1,2k-1} = c_{2k,2k}$ and $c_{2k-1,2k} = 0$, which results in T_k being the identity matrix and hence $c'_{2k-1,2k} = 0$, then from (3.21) one obtains the inequality

$$\kappa_2^2(A') > \sum_{k=1}^{r} c'^2_{2k-1,2k} \tag{3.25}$$

Assuming that $\sum_{k=1}^{r} c'^2_{2k-1,2k} \geq \frac{1}{2n} N^2(C')$, then from the fact that the Euclidean norm is invariant under orthogonal transformations and from (3.25) one obtains relation (3.17).

From Lemmas 1 and 2 it can be seen that in order to obtain the largest possible value of $\Delta N^2(A)$, the matrix A should be subjected to the orthogonal transformation $M^t A M$ where M is a permutation matrix determined as follows: Let $A'' = M^t A M$ and $C'' = A'' A''^t - A''^t A''$, then M is chosen such that each 2×2 diagonal submatrix C''_{kk} has an element $c''_{2k-1,2k}$ of at least average absolute value of all the off-diagonal elements of C" if any, and/or the difference $(c''_{2k-1,2k-1} - c''_{2k,2k})$ different from zero. For example, in order to bring the off-diagonal element c_{uv}, $(u < v)$, of maximum absolute value in the position (1,2) M is given by $I_{1u} I_{2v}$ where $I_{ij} = I - (e_i - e_j)(e_i - e_j)^t$. Essentially $I^t_{ij} A I_{ij}$ has the i-th and j-th rows and columns of A exchanged.

After the matrix A is "prepared" by the transformation M, $A' = P^t A'' P$ will produce a matrix C' whose off-diagonal elements $c'_{2k-1,2k}$ are of such magnitudes that $\sum_{k=1}^{r} c'^2_{2k-1,2k}$ is at least equal to $(1/2n) N^2(C)$.

Theorem. Let $A = A_1$ be a diagonable matrix with an even order $n = 2r$ and $N^2(A) \leq 1$. Let $A_{\ell+1} = U_\ell^{-1} A_\ell U_\ell$, where $U_\ell = M_\ell P_\ell Q_\ell$. If these transformations are defined as follows:

(i) M_ℓ is chosen as discussed above.

(ii) $P_\ell = \text{diag}\,(T_1^{(\ell)}, T_2^{(\ell)}, \ldots, T_r^{(\ell)})$

in which,

$$T_k^{(\ell)} = \begin{bmatrix} \cos \varphi_k^{(\ell)} & \sin \varphi_k^{(\ell)} \\ -\sin \varphi_k^{(\ell)} & \cos \varphi_k^{(\ell)} \end{bmatrix}$$

with,

$$\tan 2\varphi_k^{(\ell)} = \frac{c_{2k-1,2k-1}^{(\ell)} - c_{2k,2k}^{(\ell)}}{2c_{2k-1,2k}^{(\ell)}}$$

(iii) $Q_\ell = \text{diag}\,(S_1^{(\ell)}, S_2^{(\ell)}, \ldots, S_r^{(\ell)})$

in which,

$$S_1^{(\ell)} = S_2^{(\ell)} = \ldots = S_r^{(\ell)} = \begin{bmatrix} \cosh \psi_\ell & \sinh \psi_\ell \\ \sinh \psi_\ell & \cosh \psi_\ell \end{bmatrix}$$

with,

$$\tanh 4\psi_\ell = -2\kappa(A_\ell')/\kappa_1(A_\ell')$$

where,

$$A_\ell' = (M_\ell P_\ell)^t A_\ell (M_\ell P_\ell)$$

Then, $\lim_{\ell \to \infty} N^2(C_\ell) = 0$.

Proof. With no loss of generality assume that $M_\ell = I$. By Lemma 2, $\kappa_2^2(A_\ell') \geq \frac{1}{2n} N^2(C_\ell)$. From (3.3), $(D_{km}^{(\ell)})^2 + (E_{km}^{(\ell)})^2 \leq 2N^2(A_{km}^{(\ell)})$, then (3.4) yields, $\kappa_1(A_\ell) \leq 2N^2(A_\ell) \leq 2$. Since the Euclidean norm is invariant under orthogonal transformations, then $\kappa_1(A_\ell') \leq 2$, and hence by Lemma 1,

$$\Delta N^2(A_\ell) \geq \kappa_2^2(A_\ell')/\kappa_1(A_\ell') \geq \frac{1}{4n} N^2(C_\ell)$$

But since $N^2(A_\ell)$ is a decreasing monotone function bounded below by $\sum_i |\lambda_i|^2$, where λ_i are the eigenvalues of A, [10], then $\Delta N^2(A_\ell) \to 0$ as $\ell \to \infty$. Hence $N^2(C_\ell) \to 0$, and A_ℓ is arbitrarily close to being normal.

Let A be a 128 × 128 matrix. Using one quadrant of the ILLIAC IV, (64 PE's), the matrix can be stored in memory such that for a given m the 2 × 2 submatrices A_{km} (k = 1,2,...,64) are assigned to the m-th PE. Once the matrix C is determined by parallel multiplication and stored in the same way; i.e., the k-th PE contains the submatrix C_{kk}, the 64 angles φ_k can then be determined simultaneously. Also for each k the submatrices $A'_{km} = T_k^t A_{km} T_m$ are computed simultaneously for all m, hence the updated matrix $A' = P^t AP$ is computed with all the PE's working. Similarly the quantities D'_{km}, E'_{km}, and B'_{km} of the submatrices A'_{km}, and consequently the submatrices $S^{-1}A'_{km}S$ are computed with full efficiency. This part of the algorithm has been coded and successfully tested on the ILLIAC IV simulator [1].

Once the matrix A is reduced to a matrix $\tilde{A}$ which is practically normal, then for any diagonal submatrix

$$\begin{bmatrix} \tilde{a}_{pp} & \tilde{a}_{pq} \\ \tilde{a}_{qp} & \tilde{a}_{qq} \end{bmatrix}$$

either $\tilde{a}_{pq} = \tilde{a}_{qp}$; or $\tilde{a}_{pq} = -\tilde{a}_{qp}$ and $\tilde{a}_{pp} = \tilde{a}_{qq}$, to within a reasonable computational error. The matrix $\tilde{A}$ is reduced to the diagonal form by the unitary transformations $V_j^* \tilde{A}_j V_j$, (j = 1,2,3,...), where $V_j = \prod_{k=1}^{2m-1} (R_k)_j$, as in Section 2, is the transformation matrix of the j-th sweep. For each off-diagonal element $\tilde{a}_{pq}$ or $\tilde{a}_{qp}$ above the diagonal, the elements of the diagonal submatrices of R_k are given by

(a) $\tilde{a}^{(k)}_{pq} = \tilde{a}^{(k)}_{qp}$:

the elements $R^{(k)}_{pp}$, $R^{(k)}_{qq}$, $R^{(k)}_{pq}$, and $R^{(k)}_{qp}$ are determined as in Section 2.

(b) $\tilde{a}^{(k)}_{pq} = -\tilde{a}^{(k)}_{qp}$ and $\tilde{a}^{(k)}_{pp} = \tilde{a}^{(k)}_{qq}$:

$R^{(k)}_{pp} = R^{(k)}_{qq} = \frac{1}{\sqrt{2}}$; $R^{(k)}_{pq} = R^{(k)}_{qp} = \frac{i}{\sqrt{2}}$ where $i = \sqrt{-1}$, [5].

Denoting the resulting matrix by $\Lambda = Y^{-1} A Y$, the diagonal elements of Λ are then the eigenvalues of A, and the columns of the matrix $Y = (\Pi_{\ell} U_{\ell}) (\Pi_j V_j)$ are the corresponding eigenvectors.

4. Acknowledgment. The author would like to thank Professor Daniel L. Slotnick, Director of the Center of Advanced Computation, University of Illinois, for introducing him to the subject of parallel computation and for providing encouragement and guidance throughout the investigation. Special thanks go to the referee for his valuable comments and criticism of the presentation.

APPENDIX

The orthogonal matrix $R(p,q,\alpha_{pq}^{(k)})$ differs from the identity matrix by a 2×2 diagonal submatrix whose elements are,

(A.1) $R_{pp} = R_{qq} = \cos \alpha_{pq}^{(k)}$; $R_{pq} = -R_{qp} = \sin \alpha_{pq}^{(k)}$ where $p < q$. In order to eliminate the off-diagonal element $a_{pq}^{(k)}$, the angle α_{pq} is chosen such that

$$\text{(A.2)} \qquad \tan 2\alpha_{pq}^{(k)} = \frac{2a_{pq}^{(k)}}{a_{pp}^{(k)} - a_{qq}^{(k)}}$$

in which $\alpha_{pq}^{(k)}$ is restricted by $|\alpha_{pq}^{(k)}| \leq \pi/4$, [6].

Let $t_k = |2a_{pq}^{(k)}|$, $x_k = |a_{pp}^{(k)} - a_{qq}^{(k)}|$, $y_k = (t_k^2 + x_k^2)^{\frac{1}{2}}$ then,

$$\text{(A.3)} \qquad \cos^2 \alpha_{pq}^{(k)} = \tfrac{1}{2}(1 + \frac{x_k}{y_k}); \quad \sin^2 \alpha_{pq}^{(k)} = \tfrac{1}{2}(1 - \frac{x_k}{y_k})$$

Since $|\alpha_{pq}^{(k)}| \leq \pi/4$, then $\cos \alpha_{pq}^{(k)}$ will always be taken positive and $\sin \alpha_{pq}^{(k)}$ will be of the same sign as $[2a_{pq}^{(k)}/ (a_{pp}^{(k)} - a_{qq}^{(k)})]$.

REFERENCES

[1] W. Bernhard, "ILLIAC IV Codes for Jacobi and Jacobi-like Algorithms," Center for Advanced Computation Document No. 19, University of Illinois, Urbana, Illinois.

[2] R. L. Davis, "The ILLIAC IV Processing Element," IEEE Transactions on Computers, V. C-18, No. 9, Sept. 1969, pp. 800-816.

[3] P. J. Eberlein, "A Jacobi-like Method for the Automatic Computation of Eigenvalues and Eigenvectors of an Arbitrary Matrix," J. Soc. Indust. Appl. Math., V. 10, No. 1, March 1962, pp. 74-88. MR25 #2699.

[4] P. J. Eberlein and J. Boothroyd, "Solution to the Eigenproblem by a Norm Reducing Jacobi Type Method," Numerische Mathematik 11, 1968, pp. 1-12.

[5] H. H. Goldstine & L. P. Horwitz, "A Procedure for the Diagonalization of Normal Matrices," J. Assoc. Comput. Mach., V. 6, 1959, pp. 176-195. MR21 #426.

[6] P. Henrici, "On the Speed of Convergence of Cyclic and Quasicyclic Jacobi Methods for Computing Eigenvalues of Hermitian Matrices," J. Soc. Indust. Appl. Math., V. 6, No. 2, June 1958, pp. 144-162. MR20 #2084.

[7] M. Knowles, B. Okawa, Y. Muroka, and R. Wilhelmson, "Matrix Operations on ILLIAC IV," Dept. of Computer Science, University of Illinois, Urbana, Illinois, March 1967, ILLIAC IV Document No. 52.

[8] D. J. Kuck, "ILLIAC IV Software and Application Programming," IEEE Transactions on Computers, V. C-17, No. 8, August 1968, pp. 758-770.

[9] D. C. McIntyre, "An Introduction to the ILLIAC IV Computer," Datamation, April 1970, pp. 60-67.

[10] L. Mirsky, "On the Minimization of Matrix Norms," Amer. Math. Monthly, 65, 1958, pp. 106-107. MR20 #3169.

[11] A. Sameh & L. Han, "Eigenvalue Problems," Department of Computer Science, University of Illinois, Urbana, Illinois, April 1968, ILLIAC IV Document No. 127.

[12] D. L. Slotnick, et al, "The ILLIAC IV Computer," IEEE Transaction on Computers, V. C-17, No. 8, August 1968, pp. 746-757.

[13] J. H. Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press, Oxford, 1965. MR32 #1894.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	Center for Advanced Computation University of Illinois at Urbana-Champaign Urbana, Illinois 61801
2a. REPORT SECURITY CLASSIFICATION	UNCLASSIFIED
2b. GROUP	
3. REPORT TITLE	On Jacobi and Jacobi-Like Algorithms for a Parallel Computer
4. DESCRIPTIVE NOTES (Type of report and inclusive dates)	Research Report
5. AUTHOR(S) (First name, middle initial, last name)	Ahmed A. Sameh
6. REPORT DATE	December 11, 1972
7a. TOTAL NO. OF PAGES	30
7b. NO. OF REFS	13
8a. CONTRACT OR GRANT NO.	DAHC04 72-C-0001
b. PROJECT NO.	ARPA Order No. 1899
c.	
d.	
9a. ORIGINATOR'S REPORT NUMBER(S)	CAC Document No. 57
9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
10. DISTRIBUTION STATEMENT	Copies may be requested from the address given in (1) above.
11. SUPPLEMENTARY NOTES	
12. SPONSORING MILITARY ACTIVITY	U.S. Army Research Office-Durham Duke Station, Durham, North Carolina

13. ABSTRACT

Many existingalgorithms for obtaining the eigenvalues and eigenvectors of matrices would make poor use of such a powerful parallel computer as the ILLIAC IV. In this paper Jacobi's algorithm for real symmetric or complex Hermitian matrices, and a Jacobi-like algorithm for real non-symmetric matrices developed by P. J. Eberlein, are modified so as to achieve maximum efficiency for the parallel computations.

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Parallel Computers						
ILLIAC IV						
Jacobi's Algorithm						
Jacobi-like Algorithm						
Orthogonal Transformations						
Eigenvalues and Eigenvectors						
Normal Matrix						

UNCLASSIFIED

Security Classification